

# Bayesian inverse problems for Burgers and Hamilton-Jacobi equations with white noise forcing

Viet Ha Hoang

Division of Mathematical Sciences,

School of Physical and Mathematical Sciences,

Nanyang Technological University, Singapore 637371

## Abstract

The paper formulates Bayesian inverse problems for inference in a topological measure space given noisy observations. Conditions for the validity of the Bayes formula and the well-posedness of the posterior measure are studied. The abstract theory is then applied to Burgers and Hamilton-Jacobi equations on a semi-infinite time interval with forcing functions which are white noise in time. Inference is made on the white noise forcing, assuming the Wiener measure as the prior.

## 1 Introduction

Bayesian inverse problems are attracting considerable attention due to their importance in many applications ([15]). Given a functional  $\mathcal{G} : X \rightarrow \mathbb{R}^m$  where  $X$  is a probability space, the observation  $y \in \mathbb{R}^m$  of  $\mathcal{G}$  is subject to an unbiased noise  $\sigma$ :

$$y = \mathcal{G}(x) + \sigma. \quad (1.1)$$

The inverse problem determines  $x$  given the noisy observation  $y$ .

Cotter et al. [5] formulate a rigorous mathematical framework for Banach spaces  $X$ , assuming a Gaussian prior probability measure. They prove the Bayes formula for this infinite dimensional setting, and determine the Radon-Nikodym derivative of the posterior measure with respect to the prior. When  $\mathcal{G}(x)$  grows polynomially, using Fernique theorem, they show that the posterior measure continuously depends on  $y$ ; the problem is thus well-posed.

Cotter et al. [5] then apply the framework to Bayesian inverse problems for Navier Stokes equation for data assimilation in fluid mechanics. They consider Eulerian and Lagrangian data assimilation where the fluid velocity and traces of passive floats are observed respectively. They make inference on the initial velocity, and in the case of model errors, on the forcing. The random forcing function is an Ornstein-Uhlenbeck process which defines a

Gaussian probability measure on a Hilbert function space. Computationally, the posterior measure is sampled via the Markov Chain-Monte Carlo method in Cotter et al. [6].

In this paper, we are interested in Bayesian inverse problems for partial differential equations where forcing is white noise in time. We recover the white noise forcing on a semi-infinite time interval  $(-\infty, t]$ , given noisy observations. Semi-infinite time intervals are interested when the long time behaviour of the solutions is desired. We therefore depart from the Banach space framework of Cotter et al. [5] and consider the measure space setting instead. Our first aim is to formulate Bayesian inverse problems in measure spaces, for which we establish the validity of the Bayes formula and the well-posedness of the posterior measure. Our second aim is to apply the theory to inference on partial differential equations with a forcing function which is white noise in time, in particular, we study the randomly forced Burgers and Hamilton-Jacobi equations.

Burgers equation is a well known model for turbulence (see, e.g., [4]). It is simpler than Navier-Stokes equation due to the absence of the pressure but still keeps the essential nonlinear features. It also arises in many problems in non-equilibrium statistical mechanics. Burgers equation and equations of the Burgers type are employed as models for studying data assimilation in Bardos and Pironneau [2], Lundvall et al. [13], Apte et al. [1] and Freitag et al. [8]. We consider the case where spatially, the forcing is the gradient of a potential. Assuming that the velocity is also a gradient, the velocity potential satisfies a Hamilton-Jacobi equation with random forcing. This is the Kardar-Parisi-Zhang (KPZ) equation that describes lateral growth of an interface. Burgers equation with gradient forcing and Hamilton Jacobi equation are thus closely related. In the dissipative case, inverse problems to determine the dissipative coefficient for the KPZ equation was studied in Lam and Sander [12]. We will only concentrate on inviscid problems. Inviscid Burgers equation possesses many shocks where the solution is not continuous.

Burgers equation with white noise forcing in the spatially periodic case is studied by E et al. [7] in one dimension and by Iturriaga and Khanin [10] in multidimensions. They prove that almost surely, there is a unique solution that exists for all time. They establish many properties for shocks and for the dynamics of the Lagrange minimizers of the Lax-Oleinik functional. Much less is understood in the non-compact case. Hoang and Khanin [9] study the special case where the forcing potential has a large maximum and a small minimum. They prove that there is a solution that exists for all time. It is the limit of finite time solutions with the zero initial condition.

For Hamilton-Jacobi equation (3.2), we assume that the solution (i.e. the velocity potential  $\phi$ ) is observed at fixed spatial points at fixed moments. This assumption is reasonable as  $\phi$  is continuous and is well defined everywhere (though it is determined within an additive constant). Solutions for Burgers equation (3.1) are not defined at shocks, but for a given time, they are in  $L^1_{\text{loc}}$ , so observations in the form of a continuous functional in  $L^1_{\text{loc}}$  are

well-defined.

In section 2, we establish the Bayes formula for the posterior measures of Bayesian inverse problems in measure spaces, and propose conditions under which the problems are well-posed. Polynomial growth and Fernique theorem are no longer used as we do not assume a Gaussian prior. The results are generalizations of those in the Banach space setting by Cotter et al. in [5]. Section 3 introduces the randomly forced Burgers and Hamilton-Jacobi equations and formulates the Bayesian inverse problems. The spatially periodic case is considered in section 4, using the results for random Burgers and Hamilton-Jacobi equations by E et al. [7] and Iturriaga and Khanin [10]. We establish the validity of the Bayes formula and the posterior measure's local Lipschitzness with respect to the Hellinger metric; the main results are recorded in Theorems 4.6 and 4.11. The non-periodic case is studied in section 5, using the results by Hoang and Khanin [9] where we restrict our consideration to forcing potentials that satisfy Assumption 5.1, with a large maximum and a small minimum. Under this assumption, for both Hamilton-Jacobi and Burgers equations, the Bayes formula is valid and the posterior measure is locally Lipschitz with respect to the Hellinger metric. The main results are presented in Theorem 5.8 and 5.13. Section 6 proves some technical results that are used in the previous two sections.

Throughout the paper, we define  $\langle \cdot, \cdot \rangle_\Sigma = \langle \Sigma^{-1/2} \cdot, \Sigma^{-1/2} \cdot \rangle$  for a positive definite matrix  $\Sigma$ ; the corresponding norm  $|\Sigma^{-1/2} \cdot|$  is denoted as  $|\cdot|_\Sigma$ . We denote by  $c$  various positive constants that do not depend on the white noise; the value of  $c$  can differ from one appearance to the next.

## 2 Bayesian inverse problems in measure spaces

We study Bayesian inverse problems in measure spaces, in particular, the conditions for the validity of the Bayes formula, and the well-posedness of the problems. In [5], Cotter et al. require the functional  $\mathcal{G}$  in (1.1) to be measurable with respect to the prior measure for the validity of the Bayes formula. Restricting their consideration to Banach spaces, this requirement holds when  $\mathcal{G}$  is continuous. However, this is true for any topological spaces  $X$ . Cotter et al. imposed a polynomial growth on  $\mathcal{G}$  to prove the well-posedness. Using Fernique theorem, they show that when the prior measure is Gaussian, with respect to the Hellinger metric, the posterior measure is locally Lipschitz with respect to  $y$ . We will generalize this result to measure spaces, and introduce a general condition for the local Lipschitzness of the posterior measure, without imposing a Gaussian prior.

### 2.1 Problem formulation

For a topological space  $X$  with a  $\sigma$ -algebra  $\mathcal{B}(X)$ , we consider a functional  $\mathcal{G} : X \rightarrow \mathbb{R}^m$ . An observation  $y$  of  $\mathcal{G}$  is subject to a multivariate Gaussian

noise  $\sigma$ , i.e

$$y = \mathcal{G}(x) + \sigma, \quad (2.1)$$

where  $\sigma \sim \mathcal{N}(0, \Sigma)$ .

Let  $\mu_0$  be the prior probability measure on  $(X, \mathcal{B}(X))$ . Our purpose is to determine the conditional probability  $\mathbb{P}(x|y)$ . Let

$$\Phi(x, y) = \frac{1}{2}|y - \mathcal{G}(x)|_{\Sigma}^2. \quad (2.2)$$

We have the following theorem.

**Theorem 2.1** (Cotter et al. [5] Theorem 2.1) *If  $\mathcal{G} : X \rightarrow \mathbb{R}^m$  is  $\mu_0$ -measurable then the posterior measure  $\mu^y(dx) = \mathbb{P}(dx|y)$  is absolutely continuous with respect to the prior measure  $\mu_0(dx)$  and the Radon-Nikodym derivative satisfies*

$$\frac{d\mu^y}{d\mu_0}(x) \propto \exp(-\Phi(x; y)). \quad (2.3)$$

Although Cotter et al. assume that  $X$  is a Banach space, their proof holds for any topological measure spaces  $X$ .

Following [5] corollary 2.2, we have the following corollary.

**Corollary 2.2** *If  $\mathcal{G} : X \rightarrow \mathbb{R}^m$  is continuous, then the measure  $\mu^y(dx) = \mathbb{P}(dx|y)$  is absolutely continuous with respect to  $\mu_0(x)$  and the Radon-Nikodym derivative satisfies (2.3).*

*Proof* If  $\mathcal{G} : X \rightarrow \mathbb{R}^m$  is continuous, then it is measurable ([11], Lemma 1.5). The conclusion holds.  $\square$

## 2.2 Well-posedness of the problem

For Banach spaces  $X$ , when the prior measure  $\mu_0$  is Gaussian, the well-posedness of the inverse problem is established by Cotter et al. [5] by assuming that the function  $\Phi(x; y)$  has a polynomial growth in  $\|x\|_X$  fixing  $y$ , and is Lipschitz in  $y$  with the Lipschitz coefficient being bounded by a polynomial of  $\|x\|_X$ . The proof uses Fernique theorem. For a general measure space  $X$ , we make the following assumption.

**Assumption 2.3** *The function  $\Phi : X \times \mathbb{R}^m \rightarrow \mathbb{R}$  satisfies:*

- (i) *For each  $r > 0$ , there is a constant  $M(r) > 0$  and a set  $X(r) \subset X$  of positive  $\mu_0$  measure such that for all  $x \in X(r)$  and for all  $y$  such that  $|y| \leq r$*

$$0 \leq \Phi(x; y) \leq M(r).$$

- (ii) *There is a function  $G : \mathbb{R} \times X \rightarrow \mathbb{R}$  such that for each  $r > 0$ ,  $G(r, \cdot) \in L^2(X, d\mu_0)$  and if  $|y| \leq r$  and  $|y'| \leq r$  then*

$$|\Phi(x; y) - \Phi(x; y')| \leq G(r, x)|y - y'|.$$

To study the well-posedness of the Bayesian inverse problems (2.1), following Cotter et al. [5], we employ the Hellinger metric

$$d_{\text{Hell}}(\mu, \mu') = \sqrt{\frac{1}{2} \int_X \left( \sqrt{\frac{d\mu}{d\mu_0}} - \sqrt{\frac{d\mu'}{d\mu_0}} \right)^2 d\mu_0},$$

where  $\mu$  and  $\mu'$  are measures on  $(X, \mathcal{B}(X))$ . The probability measure  $\mu^y$  is determined as

$$\frac{d\mu^y}{d\mu_0} = \frac{1}{Z(y)} \exp(-\Phi(x, y)),$$

where the normalization constant is

$$Z(y) = \int_X \exp(-\Phi(x, y)) d\mu_0(x). \quad (2.4)$$

With respect to the Hellinger metric, the measure  $\mu^y$  depends continuously on  $y$  as shown in the following.

**Theorem 2.4** *Under Assumption 2.3, the measure  $\mu^y$  is locally Lipschitz in the data  $y$  with respect to the Hellinger metric: for each positive constant  $r$  there is a positive constant  $c(r)$  such that if  $|y| \leq r$  and  $|y'| \leq r$ , then*

$$d_{\text{Hell}}(\mu^y, \mu^{y'}) \leq c(r)|y - y'|.$$

*Proof* First we show that for each  $r > 0$ , there is a positive constant  $c(r)$  such that  $Z(y) \geq c(r)$  when  $|y| \leq r$ . It is obvious from (2.4) and from Assumption 2.3(i) that when  $|y| \leq r$ :

$$Z(y) \geq \mu_0(X(r)) \exp(-M(r)). \quad (2.5)$$

Using the inequality  $|\exp(-a) - \exp(-b)| \leq |a - b|$  for  $a > 0$  and  $b > 0$ , we have

$$|Z(y) - Z(y')| \leq \int_X |\Phi(x; y) - \Phi(x; y')| d\mu_0(x).$$

From Assumption 2.3(ii), when  $|y| \leq r$  and  $|y'| \leq r$ :

$$|\Phi(x; y) - \Phi(x; y')| \leq G(r, x)|y - y'|.$$

As  $G(r, x)$  is  $\mu_0$ -integrable,

$$|Z(y) - Z(y')| \leq c(r)|y - y'|. \quad (2.6)$$

The Hellinger distance satisfies

$$\begin{aligned} 2d_{\text{Hell}}(\mu^y, \mu^{y'})^2 &= \int_X \left( Z(y)^{-1/2} \exp\left(-\frac{1}{2}\Phi(x, y)\right) - \right. \\ &\quad \left. Z(y')^{-1/2} \exp\left(-\frac{1}{2}\Phi(x, y')\right) \right)^2 d\mu_0(x) \\ &\leq I_1 + I_2, \end{aligned}$$

where

$$I_1 = \frac{2}{Z(y)} \int_X \left( \exp\left(-\frac{1}{2}\Phi(x, y)\right) - \exp\left(-\frac{1}{2}\Phi(x, y')\right) \right)^2 d\mu_0(x),$$

and

$$I_2 = 2|Z(y)^{-1/2} - Z(y')^{-1/2}|^2 \int_X \exp(-\Phi(x, y')) d\mu_0(x).$$

Using again the inequality  $|\exp(-a) - \exp(-b)| \leq |a - b|$ , we deduce from (2.5)

$$\begin{aligned} I_1 &\leq c(r) \int_X |\Phi(x, y) - \Phi(x, y')|^2 d\mu_0(x) \\ &\leq c(r) \int_X (G(r, x))^2 d\mu_0(x) |y - y'|^2 \leq C(r) |y - y'|^2. \end{aligned}$$

Furthermore,

$$|Z(y)^{-1/2} - Z(y')^{-1/2}|^2 = \frac{|Z(y) - Z(y')|^2}{(Z(y)^{1/2} + Z(y')^{1/2})Z(y)Z(y')}.$$

From (2.5) and (2.6),

$$|Z(y)^{-1/2} - Z(y')^{-1/2}|^2 \leq c(r) |Z(y) - Z(y')|^2 \leq c(r) |y - y'|^2.$$

The conclusion then follows.  $\square$

### 3 Bayesian inverse problems for equations with white noise forcing

#### 3.1 Stochastic Burgers and Hamilton-Jacobi equations

We consider the inviscid Burgers equation

$$\frac{\partial u}{\partial t} + (u \cdot \nabla) u = f^W(x, t) \quad (3.1)$$

where  $u(x, t) \in \mathbb{R}^d$  is the velocity of a fluid particle in  $\mathbb{R}^d$ . The forcing function  $f^W(x, t) \in \mathbb{R}^d$  depends on  $t$  via a one dimensional white noise  $\dot{W}$  and is of the form

$$f^W(x, t) = f(x) \dot{W}(t).$$

We study the case where  $f(x)$  is a gradient, i.e. there is a potential  $F(x)$  such that

$$f(x) = -\nabla F(x).$$

Throughout this paper,  $F(x)$  is three time differentiable with bounded derivatives up to the third order. We consider the case where  $u(x, t)$  is also a gradient, i.e.  $u(x, t) = \nabla \phi(x, t)$ . The velocity potential  $\phi$  satisfies the Hamilton-Jacobi equation

$$\frac{\partial \phi(x, t)}{\partial t} + \frac{1}{2} |\nabla \phi(x, t)|^2 + F^W(x, t) = 0, \quad (3.2)$$

where the forcing potential  $F^W(x, t)$  is

$$F^W(x, t) = F(x)\dot{W}(t).$$

The viscosity solution  $\phi$  of (3.2) that corresponds to the solution  $u$  of the Burgers equation (3.1) is determined by the Lax-Oleinik variational principle. With the initial condition  $\phi_0$  at time 0,  $\phi(x, t)$  is given by

$$\phi(x, t) = \inf \left\{ \phi_0(\gamma(0)) + \int_0^t \left( \frac{1}{2} |\dot{\gamma}|^2 - F^W(\gamma(\tau), \tau) \right) d\tau \right\}, \quad (3.3)$$

where the infimum is taken among all the absolutely continuous curves  $\gamma : [0, t] \rightarrow \mathbb{R}^m$  such that  $\gamma(t) = x$ .

When the forcing function is spatially periodic, problem (3.2) possesses a unique solution (up to an additive constant) that exists for all time  $t$  (see [7, 10]). The non-compact setting is far more complicated. In ([9]), it is proved that when  $F(x)$  has a large maximum, and a small minimum, (3.2) possesses a solution that exists for all time. These imply the existence of a solution to the Burgers equation that exists for all time.

We employ the framework for measure spaces developed in section 2 to study Bayesian inverse problems for Burgers equation (3.1) and Hamilton-Jacobi equation (3.2) with white noise forcing.

### 3.2 Bayesian inverse problems

We study the Bayesian inverse problems that make inference on the white noise forcing given the observations at a finite set of times.

For Hamilton-Jacobi equation (3.2), fixing  $m$  spatial points  $x_i$  and  $m$  times  $t_i$ ,  $i = 1, \dots, m$ , the function  $\phi$  is observed at  $(x_i, t_i)$ . The observations are subject to an unbiased Gaussian noise. They are

$$z_i = \phi(x_i, t_i) + \sigma'_i,$$

where  $\sigma'_i$  are independent Gaussian random variables.

Since  $\phi$  is determined within an additive constant, we assume that it is measured at a further point  $(x_0, t_0)$  where, for simplicity only,  $t_0 < t_i$  for all  $i = 1, \dots, m$ .

Let

$$y_i = z_i - z_0 = \phi(x_i, t_i) - \phi(x_0, t_0) + \sigma'_i - \sigma'_0.$$

Let  $\sigma_i = \sigma'_i - \sigma'_0$ . We assume that  $\sigma = \{\sigma_i : i = 1, \dots, m\} \in \mathbb{R}^m$  follows a Gaussian distribution with zero mean and covariance  $\Sigma$ . The vector

$$\mathcal{G}_{HJ}(W) = \{\phi(x_i, t_i) - \phi(x_0, t_0) : i = 1, \dots, m\} \in \mathbb{R}^m \quad (3.4)$$

is uniquely determined by the Wiener process  $W$ . We recover  $W$  given

$$y = \mathcal{G}_{HJ}(W) + \sigma.$$

The solution  $u(x, t)$  of Burgers equation (3.1) is not defined at shocks. However, in the cases studied here,  $u(\cdot, t) \in L^1_{\text{loc}}(\mathbb{R}^m)$  for all time  $t$ . We assume that noisy measurements

$$y_i = l_i(u(\cdot, t_i)) + \sigma_i, \quad i = 1, \dots, m,$$

are made for  $l_i(u(\cdot, t_i))$  ( $i = 1, \dots, m$ ) where  $l_i : L^1_{\text{loc}}(\mathbb{R}^m) \rightarrow \mathbb{R}$  are continuous and bounded functionals. The noise  $\sigma = (\sigma_1, \dots, \sigma_m) \sim \mathcal{N}(0, \Sigma)$ . Letting

$$\mathcal{G}_B(W) = (l_1(u(\cdot, t_1)), \dots, l_m(u(\cdot, t_m))) \in \mathbb{R}^m, \quad (3.5)$$

given

$$y = \mathcal{G}_B(W) + \sigma,$$

we determine  $W$ .

Let  $t_{\max} = \max\{t_1, \dots, t_m\}$ . Let  $\mathcal{X}$  be the metric space of continuous functions  $W$  on  $(-\infty, t_{\max}]$  with  $W(t_{\max}) = 0$ . The space is equipped with the metric:

$$D(W, W') = \sum_{n=1}^{\infty} \frac{1}{2^n} \frac{\sup_{-n \leq t \leq t_{\max}} |W(t) - W'(t)|}{1 + \sup_{-n \leq t \leq t_{\max}} |W(t) - W'(t)|} \quad (3.6)$$

(see Stroock and Varadhan [14]). Let  $\mu_0$  be the Wiener measure on  $\mathcal{X}$ . Let  $\mu^y$  be the conditional probability defined on  $\mathcal{X}$  given  $y$ . For  $\mathcal{G}_{HJ}$  in (3.4), we define the function  $\Phi_{HJ} : \mathcal{X} \times \mathbb{R}^m \rightarrow \mathbb{R}$  by

$$\Phi_{HJ}(W; y) = \frac{1}{2} |y - \mathcal{G}_{HJ}(W)|_{\Sigma}^2, \quad (3.7)$$

and for  $\mathcal{G}_B$  in (3.5), we define

$$\Phi_B(W; y) = \frac{1}{2} |y - \mathcal{G}_B(W)|_{\Sigma}^2. \quad (3.8)$$

We will prove that the Radon-Nikodym derivative of  $\mu^y$  satisfies

$$\frac{d\mu^y}{d\mu_0} \propto \exp(-\Phi_{HJ}(W; y)),$$

and

$$\frac{d\mu^y}{d\mu_0} \propto \exp(-\Phi_B(W; y))$$

respectively, and that the posterior measure  $\mu^y$  continuously depends on  $y \in \mathbb{R}^m$ .

## 4 Periodic problems

We first consider the case where the problems are set in the torus  $\mathbb{T}^d$ . The forcing potential  $F^W(x, t)$  is spatially periodic. Randomly forced Burgers and Hamilton-Jacobi equations in a torus are studied thoroughly by E et al. in [7] for one dimension and by Iturriaga and Khanin in [10] for multidimensions. We first review their results.

## 4.1 Existence and uniqueness for periodic problems

The mean value

$$b = \int_{\mathbb{T}^d} u(x, t) dx$$

is unchanged for all  $t$ . We denote the solution  $\phi$  of the Hamilton-Jacobi equation (3.2) by  $\phi_b^W$ . It is of the form

$$\phi_b^W(x, t) = b \cdot x + \psi_b^W(x, t), \quad (4.1)$$

where  $\psi_b^W$  is spatially periodic. Let  $\mathcal{C}(s, t; \mathbb{T}^d)$  be the space of absolutely continuous functions from  $(s, t)$  to  $\mathbb{T}^d$ . For each vector  $b \in \mathbb{R}^d$ , we define the operator  $\mathcal{A}_{s,t}^{W,b} : \mathcal{C}(s, t; \mathbb{T}^d) \rightarrow \mathbb{R}$  as

$$\mathcal{A}_{s,t}^{W,b}(\gamma) = \int_s^t \left( \frac{1}{2} |\dot{\gamma}(\tau) - b|^2 - F^W(\gamma(\tau), \tau) - \frac{b^2}{2} \right) d\tau. \quad (4.2)$$

The variational formulation for the function  $\psi_b^W(x, t)$  in (4.1) is

$$\psi_b^W(x, t) = \inf_{\gamma \in \mathcal{C}(s, t; \mathbb{T}^d)} \left( \psi(\gamma(s), s) + \mathcal{A}_{s,t}^{W,b}(\gamma) \right), \quad (4.3)$$

where  $s < t$ . This is written in terms of the Lax operator as

$$\mathcal{K}_{s,t}^{W,b} \psi_b^W(\cdot, s) = \psi_b^W(\cdot, t).$$

Using integration by parts, we get

$$\begin{aligned} \mathcal{A}_{s,t}^{W,b}(\gamma) &= \int_s^t \left( \frac{1}{2} |\dot{\gamma}(\tau) - b|^2 + \nabla F(\gamma(\tau)) \cdot \dot{\gamma}(\tau) (W(\tau) - W(t)) - \frac{b^2}{2} \right) d\tau \\ &\quad - F(\gamma(s))(W(t) - W(s)). \end{aligned} \quad (4.4)$$

We study the functional  $\mathcal{A}_{s,t}^{W,b}$  via minimizers which are defined as follows ([10]).

**Definition 4.1** A curve  $\gamma : [s, t] \rightarrow \mathbb{R}^d$  such that  $\gamma(s) = x$  and  $\gamma(t) = y$  is called a minimizer over  $[s, t]$  if it minimizes the action  $\mathcal{A}_{s,t}^{W,b}$  among all the absolutely continuous curves with end points at  $x$  and  $y$  at times  $s$  and  $t$  respectively.

For a function  $\psi$ , a curve  $\gamma : [s, t] \rightarrow \mathbb{R}^d$  with  $\gamma(t) = x$  is called a  $\psi$  minimizer over  $[s, t]$  if it minimizes the action  $\psi(\gamma(s)) + \mathcal{A}_{s,t}^{W,b}(\gamma)$  among all the curves with end point at  $x$  at time  $t$ .

A curve  $\gamma : (-\infty, t] \rightarrow \mathbb{R}^d$  with  $\gamma(t) = x$  is called a one-sided minimizer if it is a minimizer over all the time intervals  $[s, t]$ .

Iturriaga and Khanin [10] proved the following theorem :

**Theorem 4.2** ([10], Theorem 1) (i) For almost all  $W$ , all  $b \in \mathbb{R}^d$  and  $T \in \mathbb{R}$  there exists a unique (up to an additive constant) function  $\phi_b^W(x, t)$ ,  $x \in \mathbb{R}^m$ ,  $t \in (-\infty, T]$  such that

$$\phi_b^W(x, t) = b \cdot x + \psi_b^W(x, t),$$

where  $\psi_b^W(\cdot, t)$  is a  $\mathbb{T}^d$  periodic function and for all  $-\infty < s < t < T$ ,

$$\psi_b^W(\cdot, t) = \mathcal{K}_{s,t}^{W,b} \psi_b^W(\cdot, s).$$

(ii) The function  $\psi_b^W(\cdot, t)$  is Lipschitz in  $\mathbb{T}^d$ . If  $x$  is a point of differentiability of  $\psi_b^W(\cdot, t)$  then there exists a unique one-sided minimizer  $\gamma_{x,t}^{W,b}$  at  $(x, t)$  and its velocity is given by the gradient of  $\phi$ :  $\dot{\gamma}_{x,t}^{W,b}(t) = \nabla \phi_b^W(x, t) = b + \nabla \psi_b^W(x, t)$ . Further, any onesided minimizer is a  $\psi_b^W$  minimizer on finite time intervals.

(iii) For almost all  $W$  and all  $b \in \mathbb{R}^d$ :

$$\lim_{s \rightarrow -\infty} \sup_{\eta \in C(\mathbb{T}^d)} \min_{C \in \mathbb{R}} \max_{x \in \mathbb{T}^d} |\mathcal{K}_{s,t}^{W,b} \eta(x) - \psi_b^W(x, t) - C| = 0.$$

(iv) The unique solution  $u_b^W(x, t)$  that exists for all time of the Burgers equation (3.1) is determined by  $u_b^W(\cdot, t) = \nabla \phi_b^W(\cdot, t)$ .

We will employ these results to study the Bayesian inverse problems formulated in section 2.

## 4.2 Bayesian inverse problem for spatially periodic Hamilton-Jacobi equation (3.2)

We first study the Bayesian inverse problem for Hamilton-Jacobi equation (3.2) in the torus  $\mathbb{T}^d$ . The observation  $\mathcal{G}_{HJ}(W)$  in (3.4) becomes

$$\mathcal{G}_{HJ}(W) = \{\psi_b^W(x_i, t_i) - \psi_b^W(x_0, t_0) : i = 1, \dots, m\}.$$

To establish the validity of the Bayesian formula (2.3), we first show that  $\mathcal{G}_{HJ}(W)$  as a map from  $\mathcal{X}$  to  $\mathbb{R}^m$  is continuous with respect to the metric  $D$  in (3.6).

The following Proposition holds.

**Proposition 4.3** The map  $\mathcal{G}_{HJ}(W) : \mathcal{X} \rightarrow \mathbb{R}^m$  is continuous with respect to the metric (3.6).

*Proof* Let  $W_k$  converge to  $W$  in the metric (3.6). There are constants  $C_k$  which do not depend on  $i$  such that

$$\lim_{k \rightarrow \infty} |\phi_b^{W_k}(x_i, t_i) - \phi_b^W(x_i, t_i) - C_k| \rightarrow 0 \quad (4.5)$$

for all  $i = 0, \dots, m$ . The proof of (4.5) is technical; we present it in section 6.1. From this we deduce

$$\lim_{k \rightarrow \infty} |(\phi_b^{W_k}(x_i, t_i) - \phi_b^{W_k}(x_0, t_0)) - (\phi_b^W(x_i, t_i) - \phi_b^W(x_0, t_0))| = 0,$$

so

$$\lim_{k \rightarrow \infty} |\mathcal{G}_{HJ}(W_k) - \mathcal{G}_{HJ}(W)| = 0.$$

The conclusion then follows.  $\square$

To establish the continuous dependence of  $\mu^y$  on  $y$ , we verify Assumption 2.3. First we prove a bound for  $\mathcal{G}_{HJ}(W)$ .

**Lemma 4.4** *There is a positive constant  $c$  which only depends on the potential  $F$  such that*

$$|\mathcal{G}_{HJ}(W)| \leq c \left( 1 + \sum_{i=1}^m \max_{t_0-1 \leq \tau \leq t_i} |W(\tau) - W(t_i)|^2 \right).$$

*Proof* Let  $\gamma$  be a  $\psi_b^W(\cdot, t_0 - 1)$  minimizer starting at  $(x_i, t_i)$ . Then

$$\psi_b^W(x_i, t_i) = \psi_b^W(\gamma(t_0 - 1), t_0 - 1) + \mathcal{A}_{t_0-1, t_i}^{W, b}(\gamma).$$

From (4.4), there is a constant  $c$  which depends only on  $b$  and  $F$  such that

$$\begin{aligned} \mathcal{A}_{t_0-1, t_i}^{W, b}(\gamma) &\geq \frac{1}{2(t_i - t_0 + 1)} \left( \int_{t_0-1}^{t_i} |\dot{\gamma}(\tau)| d\tau \right)^2 - \\ &\quad c \left( 1 + \max_{t_0-1 \leq \tau \leq t_i} |W(\tau) - W(t_i)| \right) \int_{t_0-1}^{t_i} |\dot{\gamma}(\tau)| d\tau \\ &\quad - c(1 + |W(t_i) - W(t_0 - 1)|). \end{aligned}$$

The right hand side is a quadratic form of  $\int_{t_0-1}^{t_i} |\dot{\gamma}(\tau)| d\tau$  so

$$\mathcal{A}_{t_0-1, t_i}^{W, b}(\gamma) \geq -c \left( 1 + \max_{t_0-1 \leq \tau \leq t_i} |W(\tau) - W(t_i)|^2 \right).$$

Therefore

$$\psi_b^W(x_i, t_i) \geq \psi_b^W(\gamma(t_0 - 1), t_0 - 1) - c \left( 1 + \max_{t_0-1 \leq \tau \leq t_i} |W(\tau) - W(t_i)|^2 \right).$$

Let  $\gamma'$  be the linear curve connecting  $(x_0, t_0)$  and  $(\gamma(t_0 - 1), t_0 - 1)$ . We have

$$\begin{aligned} \psi_b^W(x_0, t_0) &\leq \psi_b^W(\gamma(t_0 - 1), t_0 - 1) + \mathcal{A}_{t_0-1, t_0}^{W, b}(\gamma') \\ &\leq \psi_b^W(\gamma(t_0 - 1), t_0 - 1) + c \left( 1 + \max_{t_0-1 \leq \tau \leq t_0} |W(\tau) - W(t_0)| \right). \end{aligned}$$

Therefore

$$\psi_b^W(x_i, t_i) - \psi_b^W(x_0, t_0) \geq -c \left( 1 + \max_{t_0-1 \leq \tau \leq t_i} |W(\tau) - W(t_i)|^2 \right). \quad (4.6)$$

Let  $\bar{\gamma}$  be the linear curve connecting  $(x_0, t_0)$  and  $(x_i, t_i)$ . We then find that

$$\psi_b^W(x_i, t_i) \leq \psi_b^W(x_0, t_0) + \mathcal{A}_{t_0, t_i}^{W, b}(\bar{\gamma}).$$

Thus

$$\psi_b^W(x_i, t_i) \leq \psi_b^W(x_0, t_0) + c \left( 1 + \max_{t_0 \leq \tau \leq t_i} |W(\tau) - W(t_i)| \right). \quad (4.7)$$

From (4.6) and (4.7), we conclude that there exists a positive constant  $c$  such that

$$|\psi_b^W(x_i, t_i) - \psi_b^W(x_0, t_0)| \leq c(1 + \max_{t_0-1 \leq \tau \leq t_i} |W(\tau) - W(t_i)|^2). \quad (4.8)$$

Therefore

$$|\mathcal{G}_{HJ}(W)| \leq c \left( 1 + \sum_{i=1}^m \max_{t_0-1 \leq \tau \leq t_i} |W(\tau) - W(t_i)|^2 \right). \quad (4.9)$$

□

**Proposition 4.5** *Assumption 2.3 holds.*

*Proof* Fixing a constant  $M$ , from (4.9) there is a set of positive Wiener measure so that

$$|\mathcal{G}_{HJ}(W)| \leq M.$$

Therefore, there is a constant  $M(r)$  and a set of positive measure such that when  $|y| \leq r$

$$|\Phi_{HJ}(W; y)| = \frac{1}{2} |y - \mathcal{G}_{HJ}(W)|_\Sigma^2 \leq \frac{1}{2} \|\Sigma^{-1/2}\|_{\mathbb{R}^m, \mathbb{R}^m}^2 (|y| + M)^2 \leq M(r).$$

Assumption 2.3(i) holds.

Now we prove Assumption 2.3(ii). We have

$$\begin{aligned} |\Phi_{HJ}(W; y) - \Phi_{HJ}(W; y')| &= \frac{1}{2} \left| \langle \Sigma^{-1/2}(y - \mathcal{G}_{HJ}(W)), \Sigma^{-1/2}(y - y') \rangle + \right. \\ &\quad \left. \langle \Sigma^{-1/2}(y - y'), \Sigma^{-1/2}(y' - \mathcal{G}_{HJ}(W)) \rangle \right| \\ &\leq \frac{1}{2} \|\Sigma^{-1/2}\|_{\mathbb{R}^m, \mathbb{R}^m}^2 (|y| + |y'| + 2|\mathcal{G}_{HJ}(W)|) |y - y'|. \end{aligned} \quad (4.10)$$

Letting

$$G(r, W) = \|\Sigma^{-1/2}\|_{\mathbb{R}^m, \mathbb{R}^m}^2 \left[ r + c \left( 1 + \sum_{i=1}^m \max_{t_0-1 \leq \tau \leq t_i} |W(\tau) - W(t_i)|^2 \right) \right].$$

Obviously,  $G(r, \cdot) \in L^2(\mathcal{X}, \mu_0)$ .

From Propositions 4.3, 4.5, Corollary 2.2 and Theorem 2.4 we have:

**Theorem 4.6** *For spatially periodic Hamilton-Jacobi equation, the measure  $\mu^y(dW) = \mathbb{P}(dW|y)$  is absolutely continuous with respect to the Wiener measure  $\mu_0(dW)$ ; the Radon-Nikodym derivative satisfies*

$$\frac{d\mu^y}{d\mu_0} \propto \exp(-\Phi_{HJ}(W; y)).$$

When  $|y| \leq r$  and  $|y'| \leq r$ , there is a constant  $c(r)$  such that

$$d_{\text{Hell}}(\mu^y, \mu^{y'}) \leq c(r) |y - y'|.$$

### 4.3 Bayesian inverse problem for spatially periodic Burgers equation

We consider the Bayesian inverse problem for Burgers equation (3.1) in the torus  $\mathbb{T}^d$ . We first establish the continuity of

$$\mathcal{G}_B(W) = (l_1(u_b^W(\cdot, t_1)), \dots, l_m(u_b^W(\cdot, t_m)))$$

as a map from  $\mathcal{X}$  to  $\mathbb{R}^m$  to prove the validity of the Bayesian formula (2.3).

First we prove the uniform boundedness of  $|\dot{\gamma}(t)|$  for any minimizers  $\gamma$  on  $[t-1, t]$ . The result is proved in [7] and [10], using Gronwall's inequality. We present another proof without using this inequality, and establish explicitly the dependency of the bound on  $W$ .

**Lemma 4.7** *For any minimizer  $\gamma$  of  $\mathcal{A}_{t-1,t}^{W,b}$ ,*

$$|\dot{\gamma}(t)| \leq c(1 + \max_{t-1 \leq \tau \leq t} |W(\tau) - W(t)|^2).$$

*Proof* Let  $\bar{\gamma}$  be the linear curve connecting  $(\gamma(t), t)$  and  $(\gamma(t-1), t-1)$ . As  $|\gamma(t) - \gamma(t-1)| \leq d^{1/2}$ ,

$$\mathcal{A}_{t-1,t}^{W,b}(\bar{\gamma}) \leq c(1 + \max_{t-1 \leq \tau \leq t} |W(\tau) - W(t)|).$$

We have

$$\begin{aligned} \mathcal{A}_{t-1,t}^{W,b}(\gamma) &\geq \frac{1}{2} \left( \int_{t-1}^t |\dot{\gamma}(\tau)| d\tau \right)^2 - c(1 + \max_{t-1 \leq \tau \leq t} |W(\tau) - W(t)|) \int_{t-1}^t |\dot{\gamma}(\tau)| d\tau - \\ &\quad c(1 + |W(t-1) - W(t)|). \end{aligned}$$

Because  $\mathcal{A}_{t-1,t}^{W,b}(\gamma) \leq \mathcal{A}_{t-1,t}^{W,b}(\bar{\gamma})$ , solving the quadratic equation, we get

$$\int_{t-1}^t |\dot{\gamma}(\tau)| d\tau \leq c(1 + \max_{t-1 \leq \tau \leq t} |W(\tau) - W(t)|). \quad (4.11)$$

The minimizer  $\gamma$  satisfies (see [7] and [10])

$$\begin{aligned} \dot{\gamma}(t) &= \dot{\gamma}(s) + \int_s^t \nabla^2 F(\gamma(\tau)) \dot{\gamma}(\tau) (W(\tau) - W(t)) d\tau \\ &\quad + \nabla F(\gamma(s))(W(s) - W(t)) \end{aligned} \quad (4.12)$$

for all  $t-1 \leq s \leq t$ . Therefore

$$\begin{aligned} |\dot{\gamma}(s)| &\geq |\dot{\gamma}(t)| - c \max_{t-1 \leq \tau \leq t} |W(\tau) - W(t)| \left( 1 + \int_{t-1}^t |\dot{\gamma}(\tau)| d\tau \right) \\ &\geq |\dot{\gamma}(t)| - c(1 + \max_{t-1 \leq \tau \leq t} |W(\tau) - W(t)|^2). \end{aligned}$$

From this and (4.11), we deduce that

$$\begin{aligned} |\dot{\gamma}(t)| &\leq c(1 + \max_{t-1 \leq \tau \leq t} |W(\tau) - W(t)|^2) + c(1 + \max_{t-1 \leq \tau \leq t} |W(\tau) - W(t)|) \\ &\leq c(1 + \max_{t-1 \leq \tau \leq t} |W(\tau) - W(t)|^2). \end{aligned}$$

□

We have the following lemma on the Lipschitzness of the solution  $\phi_b^W(\cdot, t)$ .

**Lemma 4.8** *The solution  $\phi_b^W(\cdot, t)$  satisfies*

$$|\phi_b^W(x, t) - \phi_b^W(x', t)| \leq c(1 + \max_{t-1 \leq \tau \leq t} |W(\tau) - W(t)|^2) |x - x'|.$$

The lemma is proved in [10], but to show the explicit Lipschitz constant, we briefly mention the proof here.

*Proof* Let  $\gamma$  be a one-sided minimizer starting at  $(x, t)$ . We consider the curve

$$\tilde{\gamma}(\tau) = \gamma(\tau) + (x' - x)(\tau - t + 1),$$

(modulus  $\mathbb{T}^d$ ) connecting  $x'$  and  $\gamma(t-1)$ . As

$$\psi_b^W(x, t) = \psi_b^W(\gamma(t-1), t-1) + \mathcal{A}_{t-1, t}^{W, b}(\gamma),$$

and

$$\psi_b^W(x', t) \leq \phi_b^W(\gamma(t-1), t-1) + \mathcal{A}_{t-1, t}^{W, b}(\tilde{\gamma}),$$

we have

$$\begin{aligned} \psi_b^W(x', t) - \psi_b^W(x, t) &\leq \mathcal{A}_{t-1, t}^{W, b}(\tilde{\gamma}) - \mathcal{A}_{t-1, t}^{W, b}(\gamma) \\ &\leq c \left[ 1 + \max_{t-1 \leq s \leq t} |W(s) - W(t)| \left( 1 + \int_{t-1}^t |\dot{\gamma}(\tau)| d\tau \right) \right] |x - x'|. \end{aligned}$$

Similarly, the same estimate holds for  $\psi_b^W(x, t) - \psi_b^W(x', t)$ . From (4.11), we have

$$|\phi_b^W(x, t) - \phi_b^W(x', t)| \leq c(1 + \max_{t-1 \leq \tau \leq t} |W(\tau) - W(t)|^2) |x - x'|.$$

□

We now establish the continuity of  $\mathcal{G}_B(W)$  as a map from  $\mathcal{X}$  to  $\mathbb{R}^m$ .

**Proposition 4.9** *The function  $\mathcal{G}_B(W) : \mathcal{X} \rightarrow \mathbb{R}^m$  is continuous with respect to the metric  $D$  in (3.6).*

*Proof* From Lemma 4.8, when  $W_k \rightarrow W$  in the metric  $D$ ,  $\phi_b^{W_k}$  is uniformly Lipschitz in  $C^{0,1}(\mathbb{T}^d)/\mathbb{R}$ . We can therefore extract a subsequence that converges in  $C^{0,1}(\mathbb{T}^d)/\mathbb{R}$ . From (4.5), every subsequence converges to  $\phi_b^W$ . Thus  $u_b^{W_k} \rightarrow u_b^W$  whenever  $\phi_b^{W_k}$  and  $\phi_b^W$  are differentiable ( $\phi_b^{W_k}$  and  $\phi_b^W$  are differentiable almost everywhere). From Lemma 4.8,  $|u_b^{W_k}(x, t)|$  are uniformly bounded. The dominated convergence theorem implies that  $u_b^{W_k}(\cdot, t_i) \rightarrow u_b^W(\cdot, t_i)$  in  $L^1(\mathbb{T})$  for  $i = 1, \dots, m$ . □

For the well-posedness of the Bayesian inverse problem, we prove Assumption 2.3.

**Proposition 4.10** *For the spatially periodic randomly forced Burgers equation (3.1), Assumption 2.3 holds.*

*Proof* From Lemma 4.8, we deduce

$$|\mathcal{G}_B(W)| \leq c \left( 1 + \sum_{i=1}^m \max_{t_{i-1} \leq \tau \leq t_i} |W(\tau) - W(t_i)|^2 \right). \quad (4.13)$$

Therefore for each positive constant  $M$ , there exists a subset of  $\mathcal{X}$  of positive Wiener measure such that for all  $W$  in that set:  $|\mathcal{G}(W)| \leq M$ . By the same argument as in the proof of Proposition 4.5, Assumption 2.3(i) holds.

From (4.10) and (4.13), Assumption 2.3(ii) holds with

$$G(r, W) = \|\Sigma^{-1/2}\|_{\mathbb{R}^m, \mathbb{R}^m}^2 \left[ r + c \left( 1 + \sum_{i=1}^m \max_{t_{i-1} \leq \tau \leq t_i} |W(\tau) - W(t_i)|^2 \right) \right],$$

which is in  $L^2(\mathcal{X}, d\mu_0)$ .  $\square$

From Propositions 4.9 and 4.10, Corollary 2.2 and Theorem 2.4 we have

**Theorem 4.11** *For the spatially periodic randomly forced Burgers equation (3.1), the posterior measure  $\mu^y$  is absolutely continuous with respect to the prior measure  $\mu_0$  and satisfies*

$$\frac{d\mu^y}{d\mu_0} \propto \exp(-\Phi_B(W; y)),$$

where  $\Phi_B(W; y)$  is defined in (3.8), and is locally Lipschitz in  $y$  with respect to the Hellinger metric, i.e.

$$d_{\text{Hell}}(\mu^y, \mu^{y'}) \leq c(r)|y - y'|,$$

when  $|y| \leq r$  and  $|y'| \leq r$ ;  $c(r)$  depends on  $r$ .

## 5 Non-periodic problems

We consider Bayesian inverse problems making inference on white noise forcing for Burgers and Hamilton-Jacobi equations in non-compact domains. Burgers and Hamilton-Jacobi equations with white noise forcing are reasonably understood when the forcing potential  $F(x)$  has a large maximum and a small minimum (Hoang and Khanin [9]). We first present the setting up of the problem and introduce the basic results that will be used.

### 5.1 Existence and uniqueness for non-periodic problems

Let  $W(t)$  be a standard Brownian motion starting at 0 at the time 0. Let

$$E_1 = \mathbb{E}\{|W(1)|\}$$

where  $\mathbb{E}$  denotes expectation with respect to the Wiener measure. For all  $0 < t < 1$ , there is a constant  $E_2$  not depending on  $t$  such that

$$\mathbb{E}\{\max_{1-t \leq \tau \leq 1} |W(\tau) - W(1)|\} = E_2 t^{1/2}.$$

We also denote by

$$E_3 = \mathbb{E}\{\max_{0 \leq \tau \leq 1} |W(\tau) - W(1)|^2\}.$$

Following Hoang and Khanin [9], we make the following assumption.

**Assumption 5.1** (i) *The forcing potential  $F$  has a maximum at  $x^{\max}$  and a minimum at  $x^{\min}$  such that*

$$F(x^{\max}) > L \quad \text{and} \quad F(x^{\min}) < -L; \quad |x^{\max}| < a, \quad |x^{\min}| < a$$

where  $a$  and  $L$  are positive constants; there is a constant  $b > a$  such that

$$|\nabla F(x)| < K \quad \text{when } |x| < b$$

and for  $x \geq b$ ,

$$|\nabla F(x)| < K_1 \ll K \quad \text{and} \quad |F(x)| < L_1 \ll L;$$

(ii) *The function  $F(x)$  is three time differentiable;  $|\nabla^2 F(x)|$  and  $|\nabla^3 F(x)|$  are uniformly bounded for all  $x$ .*

(iii) *The constants  $L, K, L_1, K_1$  satisfy*

$$\frac{128a^4 E_2^2 K^2}{E_1^3} < L^3, \quad 8a^2 < LE_1, \quad K_1^2 < \frac{LE_1}{3E_3}, \quad \text{and} \quad L_1 < \frac{L}{16}.$$

We denote by  $B(b)$  the ball with centre 0 and radius  $b$ .

The solution of the Hamilton-Jacobi equation (3.2), denoted by  $\phi^W$  in this case, satisfies the Lax-Oleinik formula

$$\phi^W(x, t) = \inf_{\gamma \in \mathcal{C}(s, t; \mathbb{R}^d)} \{\phi(\gamma(s), s) + \mathcal{A}_{s, t}^W(\gamma)\}, \quad (5.1)$$

where  $\mathcal{C}(s, t; \mathbb{R}^d)$  is the space of absolutely continuous functions from  $(s, t)$  to  $\mathbb{R}^d$ ; the operator  $\mathcal{A}_{s, t}^W$  is defined as

$$\begin{aligned} \mathcal{A}_{s, t}^W(\gamma) = & \int_s^t \left( \frac{1}{2} |\dot{\gamma}(\tau)|^2 + \nabla F(\gamma(\tau)) \dot{\gamma}(\tau) (W(\tau) - W(t)) \right) d\tau + \\ & F(\gamma(s))(W(s) - W(t)). \end{aligned} \quad (5.2)$$

We denote (5.1) as

$$\phi^W(x, t) = \mathcal{K}_{s, t}^W \phi^W(\cdot, s).$$

Minimizers for the operator  $\mathcal{A}_{s, t}^W$  are defined as in definition 4.1.

Under Assumption 5.1, the following existence and uniqueness results for Hamilton-Jacobi equation (3.2) and Burgers equation (3.1) are established in [9].

**Theorem 5.2** (i) *With probability 1, there is a solution  $\phi^W$  to the Hamilton-Jacobi equation (3.2) such that for all  $t > 0$ ,  $\phi^W(\cdot, t) = \mathcal{K}_{s,t}^W \phi^W(\cdot, s)$  for all  $s < t$ . The solution  $\phi^W(\cdot, t)$  is locally Lipschitz for all  $t$ . Let  $\phi_s^W(x, t)$  be the solution to the Hamilton-Jacobi equation with the zero initial condition at time  $s$ , then there is a constant  $C_s$  such that*

$$\lim_{s \rightarrow -\infty} \sup_{|x| \leq M} |\phi^W(x, t) - \phi_s^W(x, t) - C_s| = 0$$

for all  $M > 0$ . The solution is unique in this sense up to an additive constant.

(ii) *For every  $(x, t)$ , there is a one-sided minimizer  $\gamma_{x,t}^W$  starting at  $(x, t)$ , and there is a sequence of time  $s_i$  that converges to  $-\infty$ , and zero minimizers  $\gamma_{s_i}$  on  $[s_i, t]$  starting at  $(x, t)$  such that  $\gamma_{s_i}$  converge uniformly to  $\gamma_{x,t}^W$  on any compact time intervals when  $s_i \rightarrow -\infty$ ; further  $\gamma_{x,t}^W$  is a  $\phi^W$  minimiser on any finite time interval. If  $x \in \mathbb{R}^m$  is a point of differentiability of  $\phi^W(\cdot, t)$  then such a one-sided minimizer starting at  $x$  at time  $t$  is unique and  $\dot{\gamma}_{x,t}^W(t) = \nabla \phi^W(x, t)$ .*

(iii) *The function  $u^W(\cdot, t) = \nabla \phi^W(\cdot, t)$  is a solution of the Burgers equation (3.1) that exists for all time. It is the limit of the solutions to (3.1) on finite time intervals with the zero initial condition. In this sense, the solution  $u^W$  is unique.*

From now on, by onesided minimizers, we mean those that are limits of zero minimizers on finite time intervals as in Theorem 5.2(ii).

The key fact that Hoang and Khanin [9] used to prove these results is that when  $t - s$  is sufficiently large, any zero minimizers starting at  $(x, t)$  over a time interval  $(s, t)$  will be inside the ball  $B(b)$  at a time in an interval  $(\bar{t}(W), t)$  where  $\bar{t}(W)$  only depends on  $W$ ,  $x$  and  $t$ .

Let

$$\alpha = \min\{1, L^2 E_1^2 / (16a^2 K^2 E_2^2)\}.$$

We define for each integer  $l$  the random variable

$$\begin{aligned} P_l = \frac{2a^2}{\alpha} + 2aK \max_{l+1-t \leq \tau \leq l+1} |W(\tau) - W(l+1)| - L|W(l) - W(l+1)| + \\ L_1|W(l) - W(l+1)| + \frac{K_1^2}{2} \max_{l \leq \tau \leq l+1} |W(\tau) - W(l+1)|^2. \end{aligned} \quad (5.3)$$

These random variables are independent and identically distributed. Their expectation satisfies

$$\mathbb{E}\{P_l\} < -\frac{LE_1}{2}, \quad (5.4)$$

which can be shown from Assumption 5.1(iii) (see [9] page 830).

Let  $t'_i$  and  $t''_i$  be the integers that satisfy

$$t_i \leq t'_i < t_i + 1 \quad \text{and} \quad t_i - 2 < t''_i \leq t_i - 1. \quad (5.5)$$

The following result is shown in [9]:

**Proposition 5.3** ([9] pages 829-831) Fixing an index  $i$ , there are constants  $c(x_i, t_i, F)$  and  $c(F)$  such that if  $\bar{t}_i(W)$  is the largest integer for which

$$\sum_{l=\bar{t}}^{t''_i-1} P_l > -c(x_i, t_i, F) \left( 1 + \sum_{j=t''_0}^{t'_i} \max_{j \leq \tau \leq j+1} |W(\tau) - W(j+1)|^2 \right) - c(F) \left( 1 + \max_{\bar{t}-2 \leq \tau \leq \bar{t}} |W(\tau) - W(\bar{t})|^2 \right), \quad (5.6)$$

for all integers  $\bar{t} \leq \bar{t}_i(W) + 2$ , then when  $s$  is sufficiently small, every zero minimizer over  $(s, t_i)$  starting at  $(x_i, t_i)$  is inside the ball  $B(b)$  at a time  $s \in [\bar{t}_i(W), t''_0]$ .

Indeed, Hoang and Khanin [9] prove the result for  $s \in [\bar{t}_i(W), t_i - 1]$  but the proof for the requirement  $s \in [\bar{t}_i(W), t''_0]$  is identical. The law of large number guarantees the existence of such a value  $\bar{t}_i(W)$ . It is obvious that all onesided minimizers in Theorem 5.2(ii) are in  $B(b)$  at a time  $s \in [\bar{t}_i(W), t''_0]$ .

We employ these results to prove the validity of the Bayes formula and the well-posedness of the Bayesian inverse problems for Burgers and Hamilton-Jacobi equations.

## 5.2 Bayesian inverse problem for Hamilton-Jacobi equation (3.2)

We consider Hamilton-Jacobi equation (3.2) with the forcing potential  $F$  satisfying Assumption 5.1. To prove the validity of the Bayes formula, we show the continuity of the function

$$\mathcal{G}_{HJ}(W) = (\phi^W(x_1, t_1) - \phi^W(x_0, t_0), \dots, \phi^W(x_m, t_m) - \phi^W(x_0, t_0)).$$

**Proposition 5.4** The function  $\mathcal{G}_{HJ}$  is continuous as a map from  $\mathcal{X}$  to  $\mathbb{R}^m$  with respect to the metric  $D$  in (3.6).

*Proof* Let  $W_k$  converge to  $W$  in the metric  $D$  of  $\mathcal{X}$ . As shown in section 6.2, there are constants  $C_k$  which do not depend on  $i$  such that

$$\lim_{k \rightarrow \infty} |\phi^{W_k}(x_i, t_i) - \phi^W(x_i, t_i) - C_k| = 0, \quad (5.7)$$

for all  $i = 1, \dots, m$ . From this

$$\lim_{k \rightarrow \infty} |(\phi^{W_k}(x_i, t_i) - \phi^{W_k}(x_0, t_0)) - (\phi^W(x_i, t_i) - \phi^W(x_0, t_0))| = 0.$$

The conclusion then follows.  $\square$

To show the continuous dependence of the posterior measure on the noisy data  $y$ , we show Assumption 2.3. We first prove the following lemmas.

**Lemma 5.5** *There is a constant  $c$  that depends on  $x_i, x_0, t_i, t_0$  and  $F$  such that*

$$\phi^W(x_i, t_i) - \phi^W(x_0, t_0) \leq c(1 + \max_{t_0 \leq \tau \leq t_i} |W(\tau) - W(t_i)|).$$

*Proof* Let  $\bar{\gamma}$  be the linear curve that connects  $(x_i, t_i)$  with  $(x_0, t_0)$ . We have

$$\phi^W(x_i, t_i) \leq \phi^W(x_0, t_0) + \mathcal{A}_{t_0, t_i}^W(\bar{\gamma}).$$

Since  $\dot{\bar{\gamma}}(\tau) = (x_i - x_0)/(t_i - t_0)$ , from (5.2)

$$\begin{aligned} |\mathcal{A}_{t_0, t_i}^W| &\leq \frac{|x_i - x_0|^2}{2|t_i - t_0|} + \max_x |\nabla F(x)| \max_{t_0 \leq \tau \leq t_i} |W(\tau) - W(t_i)| |x_i - x_0| + \\ &\quad \max_x |F(x)| |W(t_i) - W(t_0)|. \end{aligned}$$

The conclusion then follows.  $\square$ .

**Lemma 5.6** *The following estimate holds*

$$\phi^W(x_i, t_i) - \phi^W(x_0, t_0) \geq -c(x_0, x_i, F) \sum_{l=\bar{t}_i(W)}^{t'_i} (1 + \max_{l \leq \tau \leq l+1} |W(\tau) - W(l+1)|^2),$$

where the constant  $c(x_0, x_i, F)$  depends on  $x_0, x_i$  and  $F$ .

*Proof* We prove this lemma in 6.3.

From these two lemmas, we have:

**Proposition 5.7** *Assumption 2.3 holds.*

*Proof* From Lemmas 5.5 and 5.6,

$$|\mathcal{G}_{HJ}(W)| \leq S(W),$$

where

$$S(W) = c + c \sum_{i=1}^m \max_{t_0 \leq \tau \leq t_i} |W(\tau) - W(t_i)| + c \sum_{i=1}^m \sum_{l=\bar{t}_i(W)}^{t'_i} (1 + \max_{l \leq \tau \leq l+1} |W(\tau) - W(l+1)|^2),$$

where the constants do not depend on the Brownian path  $W$ .

Let  $I$  be a positive integer such that the subset  $\mathcal{X}_I$  containing all  $W \in \mathcal{X}$  that satisfy  $\bar{t}_i(W) > I$  for all  $i = 1, \dots, m$  has a positive probability. There is a constant  $\bar{c}$  such that the subset containing the paths  $W \in \mathcal{X}_I$  that satisfy

$$S(W) < \bar{c}$$

has a positive probability. Assumption 2.3(i) holds.

To prove Assumption 2.3(ii), we get from (4.10)

$$|\Phi_{HJ}(W; y) - \Phi_{HJ}(W; y')| \leq \frac{1}{2} \|\Sigma^{-1/2}\|_{\mathbb{R}^m, \mathbb{R}^m}^2 (|y| + |y'| + 2|\mathcal{G}_{HJ}(W)|) |y - y'|.$$

Let

$$G(r, W) = \|\Sigma^{-1/2}\|_{\mathbb{R}^m, \mathbb{R}^m}^2 (r + S(W)). \quad (5.8)$$

We prove that  $G(r, W)$  is in  $L^2(\mathcal{X}, \mu_0)$  in 6.4.  $\square$

From Propositions 5.4, 5.7, Corollary 2.2 and Theorem 2.4, we get

**Theorem 5.8** *Under Assumption 5.1, for Hamilton-Jacobi equation (3.2) the posterior probability measure  $\mu^y(dW) = \mathbb{P}(dW|y)$  is absolutely continuous with respect to the Wiener measure  $\mu_0(dW)$ ; its Radon-Nikodym derivative satisfies*

$$\frac{d\mu^y}{d\mu_0} \propto \exp(-\Phi_{HJ}(W; y))$$

where  $\Phi_{HJ}$  is defined in (3.7). For  $|y| \leq r$  and  $|y'| \leq r$ , there is a constant  $c(r)$  depending on  $r$  such that

$$d_{\text{Hell}}(\mu^y, \mu^{y'}) \leq c(r)|y - y'|.$$

### 5.3 Bayesian inverse problem for Burgers equation

We consider Burgers equation (3.1) in non-compact setting. We first prove that

$$\mathcal{G}_B(W) = (l_1(u(\cdot, t_1)), \dots, l_m(u(\cdot, t_m)))$$

is continuous with respect to  $W$ . To this end, we prove a bound for  $|\dot{\gamma}(t_i)|$  where  $\gamma(t_i)$  is a onesided minimizer. Another bound is proved in [9] using Gronvall's inequality. This bound involves an exponential function of the Brownian motion, which is difficult for constructing the function  $G(r, W)$  in Assumption 2.3.

**Lemma 5.9** *All one sided minimizers  $\gamma$  that start at  $(x, t_i)$  satisfy*

$$|\dot{\gamma}(t_i)| \leq c(t_i - \bar{t}_i(W))(1 + |x|^2 + \max_{\bar{t}_i(W) \leq \tau \leq t_i} |W(\tau) - W(t_i)|^2), \quad (5.9)$$

where the constant  $c$  does not depend on  $x$ ,  $t_i$  and  $\bar{t}_i(W)$ .

*Proof* Let  $s \in (\bar{t}_i(W), t_i)$  such that  $\gamma(s) \in B(b)$ . From (5.2), we have

$$\begin{aligned} \mathcal{A}_{s,t_i}^W(\gamma) &\geq \frac{1}{2(t_i - s)} \left( \int_s^{t_i} |\dot{\gamma}(\tau)| d\tau \right)^2 - c \max_{s \leq \tau \leq t_i} |W(\tau) - W(t_i)| \int_s^{t_i} |\dot{\gamma}(\tau)| d\tau \\ &\quad - c|W(s) - W(t_i)|. \end{aligned} \quad (5.10)$$

We consider the linear curve  $\bar{\gamma}$  that connects  $(x, t_i)$  and  $(\gamma(s), s)$ . As  $\dot{\bar{\gamma}}(\tau) \leq (|x| + |b|)/(t_i - s)$ ,

$$\mathcal{A}_{s,t_i}^W(\bar{\gamma}) \leq \frac{(|x| + |b|)^2}{2(t_i - s)} + c(|x| + b) \max_{s \leq \tau \leq t_i} |W(\tau) - W(t_i)| + c|W(s) - W(t_i)|. \quad (5.11)$$

The constants  $c$  in (5.10) and (5.11) do not depend on  $x$ . From  $\mathcal{A}_{s,t_i}^W(\gamma) \leq \mathcal{A}_{s,t_i}^W(\bar{\gamma})$ , solving the quadratic equation, we deduce that

$$\int_s^{t_i} |\dot{\gamma}(\tau)| d\tau \leq (t_i - s)c(1 + |x| + \max_{s \leq \tau \leq t_i} |W(\tau) - W(t_i)|),$$

where the constant  $c$  does not depend on  $x$ .

The rest of the proof is similar to that of Lemma 4.7.  $\square$

**Lemma 5.10** *For any  $x$  and  $x'$*

$$|\phi^W(x, t_i) - \phi^W(x', t_i)| \leq c(t_i - \bar{t}_i(W))^2 \left[ 1 + |x|^2 + |x'|^2 + \max_{\bar{t}_i(W) \leq \tau \leq t_i} |W(\tau) - W(t_i)|^2 \right] |x - x'|.$$

*The proof of this lemma is similar to that of Lemma 4.8.*

**Proposition 5.11** *The function  $\mathcal{G}_B$  is a continuous map from  $\mathcal{X}$  to  $\mathbb{R}^m$  with respect to the metric (3.6).*

*Proof* When  $W_k \rightarrow W$  in the metric (3.6),  $\bar{t}_i(W_k)$  as determined in Proposition 5.3 converges to  $\bar{t}_i(W)$ . From lemma 5.10,  $\phi^{W_k}(\cdot, t_i)$  are uniformly bounded in  $C^{0,1}(D)/\mathbb{R}$  for any compact set  $D \subset \mathbb{R}^m$ . Therefore, from (5.7),  $\phi^{W_k}(\cdot, t_i)$  converges to  $\phi^W(\cdot, t_i)$  in  $C^{0,1}(D)/\mathbb{R}$ . Thus,  $u^{W_k}(\cdot, t_i)$  are uniformly bounded in  $D$  and converges to  $u^W(\cdot, t_i)$  in  $L^1(D)$ . Therefore  $\mathcal{G}_B(W^k) \rightarrow \mathcal{G}_B(W)$ .  $\square$

Now we prove the continuous dependence of  $\mu^y$  on  $y$ .

**Proposition 5.12** *Assumption 2.3 holds.*

*Proof* From Lemma 5.10, we find that at the points where  $\phi^W$  is differentiable

$$|u(x, t_i)| \leq c(t_i - \bar{t}_i(W)) \left[ 1 + |x|^2 + (t_i - \bar{t}_i(W) + 1)^2 \sum_{l=\bar{t}_i(W)}^{t'_i-1} \max_{l \leq \tau \leq l+1} |W(\tau) - W(l+1)|^2 \right],$$

where  $t'_i$  is defined in (5.5). The rest of the proof is similar to that for Proposition 5.7.

The function  $G(r, W)$  is defined as

$$G(r, W) = \|\Sigma\|_{\mathbb{R}^m, \mathbb{R}^m}^2 \left[ r + c \sum_{i=1}^m (t_i - \bar{t}_i(W)) \left( 1 + |x|^2 + (t_i - \bar{t}_i(W) + 1)^2 \sum_{l=\bar{t}_i(W)}^{t'_i-1} \max_{l \leq \tau \leq l+1} |W(\tau) - W(l+1)|^2 \right) \right], \quad (5.12)$$

where  $c$  does not depend on  $x, t_i$  and  $\bar{t}_i(W)$ .

The proof that the function  $G(r, W)$  in (5.12) is in  $L^2(\mathcal{X}, d\mu_0)$  is similar to that in section 6.4 for the function  $G(r, W)$  defined in (5.8).  $\square$

From Propositions 5.11 and 5.12, Corollary 2.2 and Theorem 2.4, we deduce

**Theorem 5.13** *Under Assumption 5.1, for randomly forced Burgers equation (3.1) in non-compact setting, the posterior measure  $\mu^y$  is absolutely continuous with respect to the prior measure  $\mu_0$  and satisfies*

$$\frac{d\mu^y}{d\mu_0} \propto \exp(-\Phi_B(W; y))$$

where  $\Phi_B$  is defined in (3.8). For  $|y| \leq r$  and  $|y'| \leq r$ , there is a constant  $c(r)$  depending on  $r$  such that

$$d_{\text{Hell}}(\mu^y, \mu^{y'}) \leq c(r)|y - y'|.$$

## 6 Proofs of technical results

### 6.1 Proof of (4.5)

We use the concept of  $\varepsilon$ -narrow places defined in [10], in particular, the following result.

**Lemma 6.1** *For a sufficiently large  $T$ , if  $s_2 - s_1 = T$  and*

$$\max_{s_1 \leq \tau \leq s_2} |W(\tau) - W(s_2)| < \frac{1}{T^2},$$

*then for any minimizers  $\gamma$  on  $[s_1, s_2]$ ,*

$$|\mathcal{A}_{s_1, s_2}^{W, b}(\gamma) - \frac{b^2 T}{2}| < \varepsilon.$$

For each  $T > 0$ , for almost all  $W$ , ergodicity implies that there are infinitely many time intervals that satisfy the assumption of Lemma 6.1. The proof of Lemma 6.1 can be found in Iturriaga and Khanin [10] Lemma 12. We call such an interval  $[s_1, s_2]$  an  $\varepsilon$ -narrow place. Indeed, the definition of  $\varepsilon$ -narrow places in [10] is more general, but Lemma 6.1 is sufficient for our purpose.

To prove (4.5), we need a further result from [10].

**Lemma 6.2** ([10], equation (42)) *Let  $[s_1, s_2]$  be an  $\varepsilon$ -narrow place. Let  $\gamma$  be a  $\psi$  minimizer on  $[s_1, t]$  where  $t > s_2$  for a function  $\psi$ . Then for all  $p \in \mathbb{T}^d$ ,*

$$\psi(p) \geq \psi(\gamma(s_1)) - 2\varepsilon.$$

*Proof of (4.5)*

Let  $\gamma^{W_k}$  be a onesided minimizer starting at  $(x_i, t_i)$  for the action  $\mathcal{A}^{W_k, b}$ . As  $D(W_k, W) \rightarrow 0$  when  $k \rightarrow \infty$ , from Lemma 4.7,  $\dot{\gamma}^{W_k}(t_i)$  is uniformly bounded.

Let  $\gamma^W$  be a one sided minimizer for  $\mathcal{A}^{W, b}$ . Let  $[s_1, s_2]$  be an  $\varepsilon$ -narrow place of the action  $\mathcal{A}^{W, b}$  as defined in Lemma 6.1. When  $D(W, W_k)$  is sufficiently small,  $[s_1, s_2]$  is also an  $\varepsilon$ -narrow place for  $\mathcal{A}^{W_k, b}$ .

Let  $\gamma_1^W$  be an arbitrary  $\psi_b^W(\cdot, s_1)$  minimizer on  $[s_1, t_i]$ . From Theorem 4.2(ii),  $\gamma^W$  is a  $\psi_b^W(\cdot, s_1)$  minimizer on  $[s_1, t_i]$ , so we have from Lemma 6.2,

$$\psi_b^W(\gamma_1^W(s_1), s_1) \geq \psi_b^W(\gamma^W(s_1), s_1) - 2\varepsilon.$$

On the other hand, also from Lemma 6.2:

$$\psi_b^W(\gamma^W(s_1), s_1) \geq \psi_b^W(\gamma_1^W(s_1), s_1) - 2\varepsilon.$$

Therefore

$$|\psi_b^W(\gamma^W(s_1), s_1) - \psi_b^W(\gamma_1^W(s_1), s_1)| < 2\varepsilon.$$

Similarly, let  $\gamma_1^{W_k}$  be an arbitrary  $\psi_b^{W_k}(\cdot, s_1)$  minimizer on  $[s_1, t_i]$ . We then have

$$|\psi_b^{W_k}(\gamma^{W_k}(s_1), s_1) - \psi_b^{W_k}(\gamma_1^{W_k}(s_1), s_1)| < 2\varepsilon.$$

Fix  $\gamma_1^W$  and  $\gamma_1^{W_k}$ . Letting  $C_k = \psi_b^W(\gamma_1^W(s_1), s_1) - \psi_b^{W_k}(\gamma_1^{W_k}(s_1), s_1)$ , we deduce

$$|\psi_b^W(\gamma^W(s_1), s_1) - \psi_b^{W_k}(\gamma^{W_k}(s_1), s_1) - C_k| < 4\varepsilon. \quad (6.1)$$

We have

$$\psi_b^{W_k}(x_i, t_i) = \psi_b^{W_k}(\gamma^{W_k}(s_1), s_1) + \mathcal{A}_{s_1, s_2}^{W_k, b}(\gamma^{W_k}) + \mathcal{A}_{s_2, t_i}^{W_k, b}(\gamma^{W_k}).$$

Let  $\bar{\gamma}$  be the curve constructed as follows. For  $\tau \in [s_2, t_i]$ ,  $\bar{\gamma}(\tau) = \gamma^{W_k}(\tau)$ . On  $[s_1, s_2]$ ,  $\bar{\gamma}$  is a minimizing curve for  $\mathcal{A}_{s_1, s_2}^{W_k, b}$  that connects  $(\gamma^{W_k}(s_2), s_2)$  and  $(\gamma^W(s_1), s_1)$ . We then have

$$\begin{aligned} \psi_b^W(x_i, t_i) &\leq \psi_b^W(\gamma^W(s_1), s_1) + \mathcal{A}_{s_1, t_i}^{W_k, b}(\bar{\gamma}) \\ &= \psi_b^W(\gamma^W(s_1), s_1) + \mathcal{A}_{s_1, s_2}^{W_k, b}(\bar{\gamma}) + \mathcal{A}_{s_2, t_i}^{W_k, b}(\gamma^{W_k}). \end{aligned}$$

Therefore

$$\begin{aligned} \psi_b^W(x_i, t_i) - \psi_b^{W_k}(x_i, t_i) &\leq (\psi_b^W(\gamma^W(s_1), s_1) - \psi_b^{W_k}(\gamma^{W_k}(s_1), s_1)) + \\ &\quad (\mathcal{A}_{s_2, t_i}^{W_k, b}(\gamma^{W_k}) - \mathcal{A}_{s_2, t_i}^{W_k, b}(\gamma^{W_k})) + (\mathcal{A}_{s_1, s_2}^{W_k, b}(\bar{\gamma}) - \mathcal{A}_{s_1, s_2}^{W_k, b}(\gamma^{W_k})). \end{aligned} \quad (6.2)$$

When  $D(W_k, W) \rightarrow 0$ ,  $|\dot{\gamma}^{W_k}(t_i)|$  are uniformly bounded so from (4.12) we deduce that  $|\dot{\gamma}^{W_k}(\tau)|$  and  $|\gamma^{W_k}(\tau)|$  are uniformly bounded for all  $s_2 \leq \tau \leq t_i$ . Therefore

$$\lim_{k \rightarrow \infty} \mathcal{A}_{s_2, t_i}^{W_k, b}(\gamma^{W_k}) - \mathcal{A}_{s_2, t_i}^{W_k, b}(\gamma^{W_k}) = 0. \quad (6.3)$$

From Lemma 6.1, as  $[s_1, s_2]$  is an  $\varepsilon$ -narrow place for both  $\mathcal{A}^{W, b}$  and  $\mathcal{A}^{W_k, b}$

$$|\mathcal{A}_{s_1, s_2}^{W_k, b}(\bar{\gamma}) - \mathcal{A}_{s_1, s_2}^{W_k, b}(\gamma^{W_k})| < 2\varepsilon. \quad (6.4)$$

From (6.1), (6.2), (6.3), (6.4) we have

$$\psi_b^W(x_i, t_i) - \psi_b^{W_k}(x_i, t_i) - C_k \leq 7\varepsilon,$$

when  $k$  is sufficiently large. Similarly,

$$\psi_b^{W_k}(x_i, t_i) - \psi_b^W(x_i, t_i) + C_k \leq 7\varepsilon,$$

when  $k$  is large. From these, we deduce

$$|\psi_b^W(x_i, t_i) - \psi_b^{W_k}(x_i, t_i) - C_k| \leq 7\varepsilon,$$

when  $k$  is sufficiently large. Hence for this particular subsequence of  $\{W_k\}$  (not renumbered)

$$\lim_{k \rightarrow \infty} |\phi_b^{W_k}(x_i, t_i) - \phi_b^W(x_i, t_i) - C_k| = 0, \quad (6.5)$$

for appropriate constants  $C_k$  which do not depend on  $x_i$  and  $t_i$ . From any subsequences of  $\{W_k\}$ , we can choose such a subsequence. We then get the conclusion.  $\square$

## 6.2 Proof of (5.7)

The proof uses  $\varepsilon$  narrow places for the non-periodic setting which are defined as follows.

**Definition 6.3** *An interval  $[s_1, s_2]$  is called an  $\varepsilon$ -narrow place for the action  $\mathcal{A}^W$  defined in (5.2) if there are two compact sets  $S_1$  and  $S_2$  such that for all minimizing curves  $\gamma(\tau)$  on an interval  $[s'_1, s'_2] \supset [s_1, s_2]$  with  $|\gamma(s'_1)| \leq b$  and  $|\gamma(s'_2)| \leq b$ , then  $\gamma(s_1) \in S_1$  and  $\gamma(s_2) \in S_2$ . Furthermore, if  $\gamma$  is a minimizing curve over  $[s_1, s_2]$  connecting a point in  $S_1$  to a point in  $S_2$ , then  $|\mathcal{A}_{s_1, s_2}^W(\gamma)| < \varepsilon$ .*

In [9] Theorem 7, such an  $\varepsilon$  narrow place is constructed as follows. Assume that  $s_2 - s_1 = T$  and

$$\max_{s_1 \leq \tau \leq s_2} |W(\tau) - W(s_2)| < \frac{1}{2T}.$$

Without loss of generality, we assume that  $s_1$  and  $s_2$  are integers. Fixing a positive constant  $c$  that depends only on the function  $F$  and  $b$ , the law of large numbers implies that if two integers  $s''_1 < s_1$  and  $s''_2 > s_2$  satisfy

$$\begin{aligned} \sum_{l=s_2}^{s''_1-1} P_l + \sum_{l=s''_1}^{s_1-1} P_l &\geq -c \left[ 1 + \max_{s''_1-2 \leq \tau \leq s''_1} |W(\tau) - W(s''_1)|^2 \right] \\ &\quad -c \left[ 1 + \max_{s''_2 \leq \tau \leq s''_2+2} |W(\tau) - W(s''_2)|^2 \right], \end{aligned}$$

then  $s''_1 \geq s_1^*$  and  $s''_2 \leq s_2^*$  where  $s_1^*$  and  $s_2^*$  depend on the path  $W$  and the constant  $c$  only ( $P_l$  is defined in (5.3)). Let

$$M_1 = \sup_{s_1^*-2 \leq \tau \leq s_1} |W(\tau) - W(s_1)|, \quad M_2 = \sup_{s_2 \leq \tau \leq s_2^*+2} |W(\tau) - W(s_2)|.$$

Hoang and Khanin ([9], pages 833-835) show that there is a continuous function  $R(M_1, M_2)$  and a constant  $c$  that depends on  $F$  such that if

$$\frac{c(R(M_1, M_2)^2 + 1)}{T} < \varepsilon,$$

then  $[s_1, s_2]$  is an  $\varepsilon$ -narrow place of  $\mathcal{A}^W$ . By using ergodicity, they show that with probability 1, an  $\varepsilon$ -narrow place constructed this way always exists and  $s_2$  can be chosen arbitrarily small. When  $D(W, W')$  is sufficiently small, then it is also an  $\varepsilon$ -narrow place for  $\mathcal{A}^{W'}$ .

We then have the following lemma (indeed Hoang and Khanin [9], in the proof of their Lemma 3, show a similar result for the kicked forcing case, but the proof for white noise forcing follows the same line).

**Lemma 6.4** *If  $\gamma$  is a  $\phi(\cdot, s_1)$  minimizer on  $[s_1, t]$ , such that  $\gamma(s_1) \in S_1$  and  $\gamma(s_2) \in S_2$  then for all  $p \in S_1$*

$$\phi(p, s_1) > \phi(\gamma(s_1), s_1) - 2\varepsilon.$$

*Proof of (5.7)*

The proof is similar to that of (4.5). Assume that  $[s_1, s_2]$  is an  $\varepsilon$ -narrow place for both  $W$  and  $W_k$ . Any one sided minimizers starting at  $(x, t)$  where  $x$  is in a compact set, when  $s_2$  is sufficiently smaller than  $t$ , must intersect the ball  $B(b)$  at a time larger than  $s_2$  and at a time smaller than  $s_1$  (see the proof of Proposition 2 in [9]). Therefore  $\gamma(s_1) \in S_1$  and  $\gamma(s_2) \in S_2$ . We then proceed as in the proof of (4.5) in section 6.1  $\square$

### 6.3 Proof of Lemma 5.6

In this section, we prove Lemma 5.6.

Consider a onesided minimizer  $\gamma$  that starts at  $(x, t_i)$ . There is a time  $s \in [\bar{t}_i(W), t''_0]$  ( $\bar{t}_i(W)$  is defined in Proposition 5.3) such that  $\gamma(s) \in B(b)$ . We then have

$$\phi^W(x_i, t_i) = \phi^W(\gamma(s), s) + \mathcal{A}_{s, t_i}^W(\gamma).$$

Let  $s'$  be the integer such that  $1 \leq s' - s < 2$ . We then have

$$\mathcal{A}_{s, t_i}^W(\gamma) = \mathcal{A}_{s, s'}^W(\gamma) + \sum_{l=s'}^{t''_i-1} \mathcal{A}_{l, l+1}^W(\gamma) + \mathcal{A}_{t''_i, t_i}^W(\gamma), \quad (6.6)$$

where  $t''_i$  is defined in (5.5). From (5.2), we have

$$\begin{aligned} \mathcal{A}_{s, s'}^W(\gamma) &\geq \frac{1}{2|s' - s|} \left( \int_s^{s'} |\dot{\gamma}(\tau)| d\tau \right)^2 - \\ &\quad \max_x |\nabla F(x)| \max_{s \leq \tau \leq s'} |W(\tau) - W(s')| \int_s^{s'} |\dot{\gamma}(\tau)| d\tau - \\ &\quad \max_x |F(x)| |W(s) - W(s')|, \end{aligned}$$

which is a quadratic form for  $\int_s^{s'} |\dot{\gamma}(\tau)| d\tau$ . Therefore, there is a constant  $c$  that depends on  $F$  such that

$$\mathcal{A}_{s, s'}^W(\gamma) \geq -c(1 + \max_{s \leq \tau \leq s'} (1 + |W(\tau) - W(s')|^2)).$$

Performing similarly for other terms in (6.6), we find that there is a constant  $c$  which does not depends on  $x, x_0, t, t_0$  such that

$$\mathcal{A}_{s, t_i}^W(\gamma) \geq -c \sum_{l=s''}^{t'_i} (1 + \max_{l \leq \tau \leq l+1} |W(\tau) - W(l+1)|^2),$$

where  $t'_i$  is defined in (5.5), and  $s''$  is the integer such that  $0 \leq s - s'' < 1$ . Thus

$$\phi^W(x_i, t_i) \geq \phi^W(\gamma(s), s) - c \sum_{l=\bar{t}_i(W)}^{t'_i} (1 + \max_{l \leq \tau \leq l+1} |W(\tau) - W(l+1)|^2). \quad (6.7)$$

Now we consider the straight line  $\gamma_1$  that connects  $(x_0, t_0)$  with  $(\gamma(s), s)$ . We have

$$\phi^W(x_0, t_0) \leq \phi^W(\gamma_1(s), s) + \mathcal{A}_{s, t_0}^W(\gamma_1).$$

Let  $t_0''$  be an integer such that  $1 \leq t_0 - t_0'' < 2$ . We write:

$$\mathcal{A}_{s, t_0}^W(\gamma_1) = \mathcal{A}_{s, s'}^W(\gamma_1) + \sum_{l=s'}^{t_0''-1} \mathcal{A}_{l, l+1}^W(\gamma_1) + \mathcal{A}_{t_0'', t_0}^W(\gamma_1).$$

The velocity  $|\dot{\gamma}_1(\tau)| \leq |b - x_0|$  so there is a constant  $c(x_0)$  such that

$$\mathcal{A}_{s, t_0}^W(\gamma_1) \leq c(x_0) \sum_{l=\bar{t}_i(W)}^{t_0'} (1 + \max_{l \leq \tau \leq l+1} |W(\tau) - W(l+1)|).$$

Therefore

$$\phi^W(x_0, t_0) \leq \phi^W(\gamma(s), s) + c(x_0) \sum_{l=\bar{t}_i(W)}^{t_0'} (1 + \max_{l \leq \tau \leq l+1} |W(\tau) - W(l+1)|). \quad (6.8)$$

From equations (6.7) and (6.8), we have

$$\phi^W(x_i, t_i) - \phi^W(x_0, t_0) \geq -c \sum_{l=\bar{t}_i(W)}^{t_i'} (1 + \max_{l \leq \tau \leq l+1} |W(\tau) - W(l+1)|^2).$$

□

## 6.4 Proof that the function $G(r, W)$ in (5.8) is in $L^2(\mathcal{X}, \mu_0)$

We now show that the function  $G(r, W)$  defined in (5.8) is in  $L^2(\mathcal{X}, \mu_0)$ . We indeed show that

$$\sum_{l=\bar{t}_i(W)}^{t_i'} (1 + \max_{l \leq \tau \leq l+1} |W(\tau) - W(l+1)|^2),$$

is in  $L^2(\mathcal{X}, \mu_0)$ .

Let  $\varepsilon$  be a small positive constant. Fixing an integer  $k$ , we consider the following events:

(A):

$$\left| \sum_{l=t_i''-m}^{t_i'} \max_{l \leq \tau \leq l+1} |W(\tau) - W(l+1)|^2 - (t_i' - t_i'' + m + 1)E_3 \right| < (t_i' - t_i'' + m + 1)\varepsilon$$

for all  $m \geq k$ ,

(B):

$$\left| \sum_{l=t_i''-m}^{t_i''-1} \max_{l \leq \tau \leq l+1} |W(\tau) - W(l+1)|^2 - mE_3 \right| < m\varepsilon$$

for all  $m \geq k$ ,

(C):

$$\left| \sum_{l=t_i''-m}^{t_i''-1} P_l - m\mathbb{E}\{P_l\} \right| < m\varepsilon$$

for all  $m \geq k$ . Assume that all of these events occur. Then from (A) and (B) for  $m = k$ , we get

$$\sum_{l=t_i''}^{t_i'} \max_{l \leq \tau \leq l+1} |W(\tau) - W(j+1)|^2 \leq (t_i' - t_i'' + 1)E_3 + (t_i' - t_i'' + 2k + 1)\varepsilon;$$

using (B) for  $m = k$  and  $m = k + 1$ , we have

$$\max_{t_i''-(k+1) \leq \tau \leq t_i''-k} |W(\tau) - W(t_i''-k)|^2 \leq (2k+1)\varepsilon + E_3;$$

using (B) for  $m = k + 1$  and  $m = k + 2$  we have

$$\max_{t_i''-(k+2) \leq \tau \leq t_i''-(k+1)} |W(\tau) - W(t_i''-(k+1))|^2 \leq (2k+3)\varepsilon + E_3;$$

and using (C) for  $m = k$  we have

$$\sum_{l=t_i''-k}^{t_i''-1} P_l \leq k(\mathbb{E}\{P_l\} + \varepsilon) \leq k\left(-\frac{LE_1}{8} + \varepsilon\right),$$

(note from (5.4) that  $\mathbb{E}\{P_l\} \leq -LE_1/8$ ).

Now suppose that all the events (A), (B) and (C) hold for  $k = t_i'' - \bar{t}_i(W)$ , then from (5.6), for  $\bar{t} = \bar{t}_i(W)$ , we get

$$\begin{aligned} (t_i'' - \bar{t}_i(W))\left(-\frac{LE_1}{8} + \varepsilon\right) &\geq -c(x_i, t_i, F)\left\{1 + (t_i' - t_i'' + 1)E_3\right. \\ &\quad \left.+ [t_i' - t_i'' + 2(t_i'' - \bar{t}_i(W)) + 1]\varepsilon\right\} \\ &\quad -c(F)\left\{1 + [4(t_i'' - \bar{t}_i(W)) + 4]\varepsilon + 2E_3\right\}. \end{aligned}$$

Thus

$$\begin{aligned} (t_i'' - \bar{t}_i(W))\left[-\frac{LE_1}{8} + \varepsilon\left(1 + 2c(x_i, t_i, F) + 4c(F)\right)\right] &\geq \\ -c(x_i, t_i, F)\left[1 + (t_i' - t_i'' + 1)E_3 + (t_i' - t_i'' + 1)\varepsilon\right] \\ &\quad -c(F)(1 + 4\varepsilon + 2E_3). \end{aligned}$$

When  $\varepsilon$  is sufficiently small, this only holds when  $t_i'' - \bar{t}_i(W)$  is smaller than an integer  $N$  which does not depend on the white noise, i.e. when  $\bar{t}_i(W) \geq$

$t''_i - N$ . If (5.6) holds but  $\bar{t}_i(W) < t''_i - N$ , then at least one of the events (A), (B) and (C) must not hold. Now we denote by  $\mathcal{X}_k$  the subset of  $\mathcal{X}$  that contains all the Brownian paths  $W$  such that at least one of the events (A), (B) and (C) does not hold. For each number  $r$ , there is a constant  $C(r, \varepsilon)$  such that

$$\mathbb{P}\{\mathcal{X}_k\} \leq \frac{C(r, \varepsilon)}{k^r}, \quad (6.9)$$

(see Baum and Katz [3] Theorem 1). Therefore

$$\begin{aligned} & \mathbb{E}\left\{\left[\sum_{l=\bar{t}_i(W)}^{t'_i} \left(1 + \max_{l \leq \tau \leq l+1} |W(\tau) - W(l+1)|^2\right)\right]^2\right\} \leq \\ & \mathbb{E}\left\{\left[\sum_{l=t''_i-N}^{t'_i} \left(1 + \max_{l \leq \tau \leq l+1} |W(\tau) - W(l+1)|^2\right)\right]^2\right\} + \\ & \sum_{k=1}^{\infty} \mathbb{E}\left\{\left[\sum_{l=t''_i-k}^{t'_i} \left(1 + \max_{l \leq \tau \leq l+1} |W(\tau) - W(l+1)|^2\right)\right]^2 \mathbf{1}_{\mathcal{X}_k}\right\}. \end{aligned}$$

From (6.9), we have

$$\begin{aligned} & \mathbb{E}\left\{\left[\sum_{l=t''_i-k}^{t'_i} \left(1 + \max_{l \leq \tau \leq l+1} |W(\tau) - W(l+1)|^2\right)\right]^2 \mathbf{1}_{\mathcal{X}_k}\right\} \\ & \leq \frac{C(r, \varepsilon)}{k^{r/2}} \left[ \mathbb{E}\left\{\left[\sum_{l=t''_i-k}^{t'_i} \left(1 + \max_{l \leq \tau \leq l+1} |W(\tau) - W(l+1)|^2\right)\right]^4\right\} \right]^{1/2} \\ & \leq \frac{C(r, \varepsilon)}{k^{r/2}} (t'_i - t''_i + k + 1)^{3/2} \left[ \mathbb{E}\left\{\left[\sum_{l=t''_i-k}^{t'_i} \left(1 + \max_{l \leq \tau \leq l+1} |W(\tau) - W(l+1)|^2\right)^4\right]\right\} \right]^{1/2} \\ & \leq C(r, \varepsilon) k^{-r/2} (t'_i - t''_i + k + 1)^2. \end{aligned}$$

When  $r$  is sufficiently large,  $\sum_{k=1}^{\infty} k^{-r/2} (t'_i - t''_i + k + 1)^2$  is finite. The assertion follows.  $\square$

**Acknowledgement** The author thanks Andrew Stuart for suggesting the metric space framework and for helpful discussions. He also thanks Kostya Khanin for introducing him to Burgers turbulence.

## References

- [1] A. Apte, D. Auroux, and M. Ramaswamy. Variational data assimilation for discrete Burgers equation. *Electron. J. Diff. Eqns., Conference*, 19:15–30, 2010.

- [2] C. Bardos and O. Pironneau. Data assimilation for conservation laws. *Methods Appl. Anal.*, 12(2):103–134, 2005.
- [3] L. E. Baum and M. Katz. Convergence rates in the law of large numbers. *Bull. Amer. Math. Soc.*, 69:771–772, 1963.
- [4] J. Bec and K. Khanin. Burgers turbulence. *Phys. Rep.*, 447(1-2):1–66, 2007.
- [5] S. L. Cotter, M. Dashti, J. C. Robinson, and A. M. Stuart. Bayesian inverse problems for functions and applications to fluid mechanics. *Inverse Problems*, 25:115008, 2009.
- [6] S. L. Cotter, M. Dashti, and A. M. Stuart. Variational data assimilation using targetted random walks. *Int. J. Numer. Meth. Fluids*, 2011 (to appear).
- [7] Weinan E, K. Khanin, A. Mazel, and Ya. Sinai. Invariant measures for Burgers equation with stochastic forcing. *Ann. of Math. (2)*, 151(3):877–960, 2000.
- [8] M. Freitag, N. K. Nichols, and C. J. Budd.  $L_1$ -regularisation for ill-posed problems in variational data assimilation. *Proc. Appl. Math. Mech.*, 10:665–668, 2010.
- [9] V. H. Hoang and K. Khanin. Random Burgers equation and Lagrangian systems in non-compact domains. *Nonlinearity*, 16(3):819–842, 2003.
- [10] R. Iturriaga and K. Khanin. Burgers turbulence and random Lagrangian systems. *Comm. Math. Phys.*, 232(3):377–428, 2003.
- [11] O. Kallenberg. *Foundations of modern probability*. Probability and its Applications (New York). Springer-Verlag, New York, second edition, 2002.
- [12] C. H. Lam and L. M. Sander. Inverse methods for interface problems. *Phys. Rev. Lett.*, 71:561–564, 1993.
- [13] J. Lundvall, V. Kozlov, and P. Wernerfelt. Iterative methods for data assimilation for Burgers' equation. *J. Inverse Ill-Posed Probl.*, 14(5):505–535, 2006.
- [14] D. W. Stroock and S. R. Srinivasa Varadhan. *Multidimensional diffusion processes*. Classics in Mathematics. Springer-Verlag, Berlin, 2006. Reprint of the 1997 edition.
- [15] A. M. Stuart. Inverse problems: A Bayesian perspective. *Acta Numerica*, 19:451–559, 2010.